

**Prioritätsbescheinigung über die Einreichung
einer Patentanmeldung****Aktenzeichen:**

10 2004 005 998.5

Anmeldetag:

06. Februar 2004

Anmelder/Inhaber:

Dr. Dietmar Ruwisch, 12557 Berlin/DE

Bezeichnung:Verfahren und Vorrichtung zur Separierung
von Schallsignalen**IPC:**

G 10 L 21/02



Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 3. Februar 2005
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag


Hoß

Verfahren und Vorrichtung zur Separierung von Schallsignalen

Die vorliegende Erfindung betrifft ein Verfahren und eine Vorrichtung zur Separierung von Schallsignalen.

5

Die Erfindung liegt auf dem Gebiet der digitalen Signalverarbeitung zum Entmischen verschiedener akustischer Signale aus unterschiedlichen Raumrichtungen, welche mit zwei Mikrofonen in bekanntem Abstand stereo aufgenommen werden.

10

Das Gebiet der Quellentrennung, auch „beam forming“ genannt, erfährt wachsende Bedeutung durch die Zunahme der mobilen Kommunikation sowie der automatischen Verarbeitung menschlicher Sprache. In sehr vielen Anwendungen tritt das Problem auf, dass das gewünschte Sprachsignal (Nutzsignal) durch verschiedene Störeinflüsse beeinträchtigt ist. Hier sind hauptsächlich Störungen durch Hintergrundgeräusche, Störungen durch andere Sprecher sowie Störungen durch Lautsprecherabgaben von Musik oder Sprache zu nennen. Die verschiedenen Störeinflüsse erfordern je nach ihrer Art und nach der Vorkenntnis über das Nutzsignal unterschiedliche Behandlungen.

20

Beispielhafte Anwendungen der Erfindung finden sich also in Kommunikationseinrichtungen, in denen die Position eines Sprechers bekannt ist, und in denen Störungen durch Hintergrundgeräusche oder andere Sprecher sowie Lautsprecherabgaben vorhanden sind. Anwendungsbeispiele sind Kfz-Freisprecheinrichtungen, in denen die Mikrofone z.B. im Rückspiegel untergebracht sind und eine sogenannte Richthyperbel auf den Fahrer gerichtet wird. In dieser Anwendung kann eine zweite Richthyperbel auf den Beifahrer gerichtet werden, so dass während eines Telefongesprächs gezielt zwischen Fahrer und Beifahrer hin- und hergeschaltet werden kann.

30

In Fällen, in denen die geometrische Position der Nutzsignalquelle zu den aufnehmenden Mikrofonen bekannt ist, ist die geometrische Quellenseparation ein mächtiges Werkzeug. Das Standardverfahren dieser Klasse von „beam forming“-

Algorithmen ist das sog. „shift and add“ Verfahren, bei welchem auf eines der Mikrofonssignale ein Filter angewendet wird, und das gefilterte Signal sodann zum zweiten Mikrofonssignal hinzuaddiert wird (siehe z.B. Haddad und Benoit, „Capabilities of a beamforming technique for acoustic measurements inside a moving car“, The 2002 International Congress and Exposition On Noise Control Engineering, Dearborn, Mi, USA, August 19-21, 2002).

Eine Erweiterung dieses Verfahrens beschäftigt sich mit „adaptiven beam forming“ bzw. „adaptiver Quellenseparation“, wo die Lage der Quellen im Raum a priori unbekannt ist und durch die Algorithmen erst ermittelt werden muss (WO 02/0671732, US6,654,719). Hier ist es das Ziel, die Lage der Quellen im Raum aus den Mikrofonssignalen zu bestimmen und nicht, wie beim „geometrischen“ beam forming, fest vorzugeben. Adaptive Verfahren erweisen sich zwar als nützlich, allerdings ist auch hier gewöhnlich a-priori-Information erforderlich, da ein Algorithmus in der Regel nicht entscheiden kann, welche der detektierten Sprachquellen Nutz- und welche Störsignal ist. Nachteilig bei allen bekannten adaptiven Verfahren ist die Tatsache, dass die Algorithmen eine gewisse Adaptionszeit benötigen, bevor ausreichende Konvergenz besteht und die Quellentrennung gelingt. Außerdem sind adaptive Verfahren prinzipiell anfälliger für diffuse Hintergrundstörungen, da diese die Konvergenz erheblich beeinträchtigen können. Ein gravierender Nachteil beim klassischen „shift and add“-Verfahren ist die Tatsache, dass sich mit zwei Mikrofonen lediglich zwei Signalquellen voneinander separieren lassen und die Dämpfung von diffusem Hintergrundschall in der Regel nicht in ausreichendem Maße gelingt.

25

Es ist daher Aufgabe der Erfindung, ein Verfahren und eine Vorrichtung zur Separierung von Schallsignalen anzugeben, das mit nur zwei Mikrofonen mehr als zwei Schallquellen im Raum separieren kann, ohne die für adaptive Verfahren typische Anfälligkeit für diffuse Hintergrundstörungen aufzuweisen.

30

Erfindungsgemäß wird diese Aufgabe durch ein Verfahren nach Anspruch 1 bzw. eine Vorrichtung nach Anspruch 8 gelöst. Vorteilhafte Weiterbildungen der Erfindung sind in den jeweiligen Unteransprüchen definiert.

Das erfindungsgemäße Verfahren benötigt keine Konvergenzzeit und kann mit zwei Mikrofonen mehr als zwei Schallquellen im Raum separieren, sofern diese in ausreichendem Maße räumlich getrennt sind. Das Verfahren stellt nur
5 geringe Anforderungen an Speicherbedarf und Rechenleistung, und es ist sehr stabil gegenüber diffusen Störsignalen. Anders als beim herkömmlichen beam forming lassen sich solche diffusen Störungen effektiv dämpfen. Wie bei allen Zwei-Mikrofon-Verfahren sind die Raumbereiche, zwischen denen das Verfahren differenzieren kann, rotationssymmetrisch zur Mikrofon-Achse, d.h. zu der Geraden, welche durch
10 die beiden Mikrofonpositionen definiert ist. In einem Schnitt durch den Raum, welcher die Symmetrieachse enthält, entspricht der Raumbereich, in dem sich eine Schallquelle befinden muss um als Nutzsignal betrachtet zu werden, einer Hyperbel. Der Winkel ϑ_0 , den der Scheitel der Hyperbel zur Symmetrieachse einnimmt, ist frei wählbar, und die Breite der Hyperbel, welche durch einen Winkel γ_{3db} bestimmt wird,
15 ist ebenfalls ein wählbarer Parameter. Mit nur zwei Mikrofonen lassen sich gleichzeitig Ausgangssignale zu beliebigen, verschiedenen Winkeln ϑ_0 erzeugen, wobei die Trennschärfe zwischen den Bereichen mit dem Überlappingsgrad der entsprechenden Hyperbeln abnimmt. Schallquellen innerhalb einer Hyperbel werden als Nutzsignale betrachtet und mit weniger als 3 db gedämpft. Störsignale werden in
20 Abhängigkeit ihres Einfallswinkels ϑ eliminiert, wobei eine Dämpfung von >25db für Einfallswinkel ϑ außerhalb der Akzeptanzhyperbel erreichbar ist.

Das Verfahren arbeitet im Frequenzbereich. Das einer Richthyperbel zuzuordnende Signalspektrum entsteht durch Multiplikation einer Korrekturfunktion
25 $K_2(x_1)$ sowie einer Filterfunktion $F(f,T)$ mit dem Signalspektrum $M(f,T)$ eines der Mikrofone. Die Filterfunktion entsteht durch spektrale Glättung (z.B. durch Diffusion) einer Zuordnungsfunktion $Z(\vartheta-\vartheta_0)$, wobei im Argument der Zuordnungsfunktion der berechnete Einfallswinkel ϑ einer spektralen Signalkomponente steht. Dieser Einfallswinkel ϑ wird aus dem Phasenwinkel φ des komplexen Quotienten der
30 Spektren der beiden Mikrofonensignale, $M_2(f,T)/M_1(f,T)$, ermittelt, indem man φ mit der Schallgeschwindigkeit c multipliziert und durch $2\pi fd$ dividiert, wobei d den Mikrofonabstand bezeichnet. Das Ergebnis $x_1 = \varphi c / 2\pi fd$, welches zugleich das

Argument der Korrekturfunktion $K_2(x_1)$ ist, liefert nach Beschränkung $x=K_1(x_1)$ auf einen Betrag kleiner oder gleich eins den Kosinus des Einfallswinkels ϑ , welcher im Argument der Zuordnungsfunktion $Z(\vartheta-\vartheta_0)$ steht; $K_1(x_1)$ bezeichnet dabei eine weitere Korrekturfunktion.

5

Kurzbeschreibung der Abbildungen:

Fig. 1 zeigt die Definition des Einfallswinkels ϑ durch die Positionen beiden Mikrofone, deren Signale verarbeitet werden.

10

Fig. 2 zeigt beispielhaft eine Zuordnungsfunktion $Z(\vartheta)$ mit Halbwertsbreite $2\gamma_{3db}$, aus welcher eine Hyperbel mit Scheitel bei $\vartheta=0$ resultiert.

Fig. 3 zeigt eine Hyperbel mit Scheitel bei $\vartheta=\vartheta_0$, die die Richtcharakteristik der Quellenseparation bestimmt. Signale innerhalb des durch die Hyperbel definierten Raumbereichs werden mit einer Dämpfung $<3db$ als Nutzsignal ausgegeben

Fig. 4 zeigt den Aufbau des Quellenseparators, in welchem die Zeitsignale zweier Mikrofone, $m_1(t)$ und $m_2(t)$, in einer Stereo-Abtast- und -Fourier-Transformator-Einheit (20) zu Spektren $M_1(f,T)$ und $M_2(f,T)$ transformiert werden, wobei T den Zeitpunkt der Entstehung der Spektren bezeichnet. Aus den Spektren wird in der ϑ -Berechnungseinheit (30) der frequenzabhängige Einfallswinkel $\vartheta(f,T)$ sowie das korrigierte Mikrofonspektrum $M(f,T)$ berechnet, woraus in Signalgeneratoren (40) für verschiedene Richtwinkel ϑ_0 Ausgangssignale $s_{\vartheta_0}(t)$ entstehen.

Fig. 5 zeigt den Aufbau der ϑ -Berechnungseinheit (30), in welcher der Phasenwinkel $\phi(f,T)$ einer spektralen Komponenten des komplexen Quotienten der beiden Mikrofonspektren $M_1(f,T)$ und $M_2(f,T)$ berechnet wird, welcher sodann mit der Schallgeschwindigkeit c zu multiplizieren und durch $2\pi fd$ zu dividieren ist, wobei d den Mikrofonabstand bezeichnet. Bei dieser Operation entsteht die Größe $x_1(f,T)$,

30

welche das Argument der beiden Korrekturfunktionen K_2 und K_1 darstellt. Mit diesen Korrekturfunktionen entsteht das korrigierte Mikrofonspektrum $M(f,T)=M_1(f,T)*K_2(x_1(f,T))$ sowie die Größe $x(f,T)=K_1(x_1(f,T))$, aus welcher durch Anwendung der Arcuskosinus-Funktion der Einfallswinkel $\vartheta(f,T)$ zu berechnen ist.

5

Fig. 6 zeigt einen Signalgenerator, in welcher eine Zuordnungsfunktion $Z(\vartheta-\vartheta_0)$ mit einem einstellbaren Winkel ϑ_0 durch spektrale Diffusion zu einer Filterfunktion $F(f,T)$ geglättet wird, welche mit dem korrigierten Mikrofonspektrum $M(f,T)$ zu multiplizieren ist. Daraus resultiert ein Ausgangsspektrum $S_{\vartheta_0}(f,T)$, aus welchem durch inverse Fouriertransformation ein Ausgangssignal $s_{\vartheta_0}(t)$ entsteht, welches die Schallsignale innerhalb des durch die Zuordnungsfunktion Z und den Winkel ϑ_0 festgelegten Raumbereichs enthält.

15

Fig. 7 zeigt exemplarisch die beiden Korrekturfunktionen $K_2(x_1)$ und $K_1(x_1)$.

20



25

30

Ein Grundgedanke der Erfindung ist es, jeder spektralen Komponente des einfallenden Signals zu jedem Zeitpunkt T einen Einfallswinkel ϑ zuzuordnen und allein anhand des berechneten Einfallswinkels zu entscheiden, ob die entsprechende Schallquelle innerhalb einer gewünschten Richthyperbel liegt, oder nicht. Um die Zugehörigkeitsentscheidung etwas abzumildern, wird anstatt einer harten Ja/Nein-Entscheidung eine „weiche“ Zuordnungsfunktion $Z(\vartheta)$ (Fig. 2) benutzt, die einen kontinuierlichen Übergang zwischen erwünschten und unerwünschten Einfallsrichtungen erlaubt, was sich vorteilhaft auf die Integrität der Signale auswirkt. Die Breite der Zuordnungsfunktion entspricht dann der Breite der Richthyperbel (Fig. 3). Durch Division der komplexen Spektren der beiden Mikrofonsignale wird zunächst für jede Frequenz f zu einem Zeitpunkt T die Phasendifferenz φ berechnet. Mit Hilfe der Schallgeschwindigkeit c und der Frequenz f der entsprechenden Signalkomponente lässt sich aus der Phasendifferenz ein Wegunterschied berechnen, der zwischen den beiden Mikrofonen liegt, wenn das Signal von einer Punktquelle ausgesandt wurde. Ist der Mikrofonabstand d bekannt, ergibt eine einfache geometrische Überlegung, dass der Quotient x_1 aus Wegunterschied und

Mikrofonabstand dem Kosinus des gesuchten Einfallswinkels entspricht. In der Praxis ist aufgrund von Störungen wie diffusem Störschall oder Raumhall die Annahme einer Punktquelle selten erfüllt, weshalb x_1 gewöhnlich nicht auf den erwarteten Wertebereich $[-1,1]$ beschränkt ist. Bevor der Einfallswinkel ϑ berechnet werden kann, ist daher noch eine Korrektur erforderlich, die x_1 auf das genannte Intervall beschränkt. Wurde zum Zeitpunkt T für jede Frequenz f der Einfallswinkel $\vartheta(f,T)$ bestimmt, ergibt sich das Spektrum des gewünschten Signals innerhalb einer Richthyperbel mit Scheitel beim Winkel $\vartheta = \vartheta_0$ durch einfache frequenzweise Multiplikation mit dem Spektrum eines der Mikrofone, also $M_1(f,T)K(\vartheta(f,T) - \vartheta_0)$. Unter Umständen ist es vorteilhaft, $K(\vartheta(f,T) - \vartheta_0)$ vor Ausführung der Multiplikation spektral zu glätten. Eine Glättung, deren Ergebnis als $F_{\vartheta_0}(f,T)$ bezeichnet sei, erhält man z.B. durch Anwendung eines Diffusionsoperators. In Fällen, in denen durch Störeinflüsse die Größe x , die zur Berechnung des Einfallswinkels dient, außerhalb ihres Wertebereichs liegt, ist es vorteilhaft, die entsprechende spektrale Komponente des Mikrofonsignals abzuschwächen, da zu vermuten ist, dass sich Störsignale überlagert haben. Dies geschieht z.B. durch Anwendung einer Korrekturfunktion, deren Argument die Größe x_1 ist. Sei $M(f,T)$ das korrigierte Mikrofonsignal, dann schreibt sich die Erzeugung des gewünschten Signalspektrums inklusive spektraler Glättung und Korrektur als $S_{\vartheta_0}(f,T) = F_{\vartheta_0}(f,T)M(f,T)$. Aus $S_{\vartheta_0}(f,T)$ entsteht durch inverse Fouriertransformation das Zeitsignal $s_{\vartheta_0}(t)$ für die entsprechende Richthyperbel mit Scheitelwinkel ϑ_0 .

Anders ausgedrückt ist es eine Grundidee der Erfindung, verschiedene Schallquellen, beispielsweise den Fahrer und den Beifahrer in einem Kraftfahrzeug, räumlich voneinander zu unterscheiden und somit beispielsweise das Nutz-Sprachsignal des Fahrers vom Stör-Sprachsignal des Beifahrers zu separieren, indem man die Tatsache ausnutzt, dass diese beiden Sprachsignale, also Schallsignale, in der Regel auch bei unterschiedlichen Frequenzen vorliegen. Die erfindungsgemäß vorgesehene Frequenzanalyse erlaubt also zunächst, das Gesamt-Schallsignal in die zwei Einzel-Schallsignale (nämlich vom Fahrer und vom Beifahrer) aufzuspalten. Es muß dann "nur noch" mit Hilfe geometrischer Überlegungen anhand der jeweiligen Frequenz jedes der beiden Schallsignale und

der zu ermittelnden Phasendifferenz zwischen dem Ausgangssignal des Mikrofons 1 und des Mikrofons 2, die jeweils diesem Schallsignal zugeordnet sind, die Einfallsrichtung jedes der beiden Schallsignale berechnet werden. Da die Geometrie zwischen beispielsweise der Position des Fahrers, der Position des Beifahrers und der Position der Mikrofone, etwa in einer Freisprecheinrichtung im Kraftfahrzeug, bekannt ist, kann dann das weiter zu verarbeitende Nutz-Schallsignal aufgrund seines anderen Einfallswinkels vom Stör-Schallsignal separiert werden.

Es folgt ein detailliertes Ausführungsbeispiel der Erfindung, das anhand der Abbildungen beschrieben wird.

Die Zeitsignale $m_1(t)$ und $m_2(t)$ zweier Mikrofone, die einen festen Abstand d zueinander haben, werden einem Rechenwerk (10) zugeführt (Fig. 4), wo sie in einer Stereo-Abtast- und -Fourier-Transformator-Einheit (20) mit einer Abtastrate f_A diskretisiert und digitalisiert werden. Eine Folge von a Abtastwerten jeweils eines der Mikrofonsignale $m_1(t)$ und $m_2(t)$ wird durch Fourier-Transformation zum komplexwertigen Spektrum $M_1(f,T)$ bzw. $M_2(f,T)$ transformiert, wobei f die Frequenz der jeweiligen Signalkomponente bezeichnet, und T den Zeitpunkt der Entstehung eines Spektrums angibt. Für die praktische Anwendung ist folgende Parameterwahl geeignet: $f_A = 11025$ Hz, $a=256$, $T = a/2 = t$. Wenn Rechenleitung und Speicherplatz es erlauben, ist jedoch $a=1024$ zu bevorzugen. Der Mikrofonabstand d sollte kleiner sein als die halbe Wellenlänge der höchsten zu verarbeitenden Frequenz, welche sich aus der Abtastfrequenz ergibt, d.h. $d < c/4f_A$. Für die oben angegebene Parameterwahl eignet sich ein Mikrofonabstand $d = 20$ mm.

25

Die Spektren $M_1(f,T)$ und $M_2(f,T)$ werden einer ϑ -Berechnungseinheit mit Spektrum-Korrektur (30) zugeführt, die aus den Spektren $M_1(f,T)$ und $M_2(f,T)$ einen Einfallswinkel $\vartheta(f,T)$ berechnet, der angibt, aus welcher Richtung relativ zur Mikrofonachse eine Signalkomponente mit Frequenz f zum Zeitpunkt T in die Mikrofone einfällt (Fig.1). Dazu wird $M_2(f,T)$ durch $M_1(f,T)$ komplex dividiert. $\phi(f,T)$ bezeichne den Phasenwinkel dieses Quotienten. Wo Verwechslungen ausgeschlossen sind, wird im folgenden das Argument (f,T) der zeit- und frequenzabhängigen Größen fortgelassen. Die genaue Rechenvorschrift zur

30

Bestimmung von φ lautet gemäß der Eulerschen Formel und den Rechenregeln für komplexe Zahlen:

$$\varphi = \arctan((\text{Re1} \cdot \text{Im2} - \text{Im1} \cdot \text{Re2}) / (\text{Re1} \cdot \text{Re2} + \text{Im1} \cdot \text{Im2})),$$

5

wobei Re1 und Re2 die Realteile und Im1 und Im2 die Imaginärteile von $M1$ bzw. $M2$ bezeichnen. Die Größe $x1 = \varphi c / 2\pi f d$ entsteht mit Hilfe der Schallgeschwindigkeit c aus dem Winkel φ , auch $x1$ ist frequenz- und zeitabhängig: $x1 = x1(f, T)$. Der Wertebereich von $x1$ muss in der Praxis mit Hilfe einer Korrekturfunktion $x = K1(x1)$ (Fig. 7) auf das
10 Intervall $[-1, 1]$ beschränkt werden. Auf die so berechnete Größe x wird durch Anwendung der Arcuskosinus-Funktion ein Einfallswinkel ϑ der betrachteten Signalkomponente errechnet, welcher von der Mikrofonachse zu messen ist, d.h. von der durch die Positionen der beiden Mikrofone definierten Geraden (Fig. 1). Unter Berücksichtigung aller Abhängigkeiten lautet damit der Einfallswinkel einer
15 Signalkomponente mit Frequenz f zum Zeitpunkt T : $\vartheta(f, t) = \arccos(x(f, T))$. Des weiteren wird mit Hilfe einer zweiten Korrekturfunktion $K2(x1)$ das Mikrofonspektrum korrigiert (Fig. 7): $M(f, T) = K2(x1)M1(f, T)$. Diese Korrektur dient dazu, in Fällen, in denen die erste Korrekturfunktion greift, die entsprechende Signalkomponente zu reduzieren, da zu vermuten ist, dass sich Störungen überlagert haben, die das Signal
20 verfälschen. Die zweite Korrektur ist optional, alternativ kann auch $M(f, T) = M1(f, T)$ gewählt werden; $M(f, T) = M2(f, T)$ ist ebenfalls möglich.

Das Spektrum $M(f, T)$ wird zusammen mit dem Winkel $\vartheta(f, T)$ einem oder mehreren Signalgeneratoren (40) zugeführt, wo mit Hilfe einer Zuordnungsfunktion
25 $Z(\vartheta)$ (Fig. 2) und einem wählbaren Winkel ϑ_0 jeweils ein auszugebendes Signal $s_{\vartheta_0}(t)$ entsteht. Dies geschieht, indem zu einem Zeitpunkt T jede spektrale Komponente des Spektrums $M(f, T)$ mit der entsprechenden Komponente eines ϑ_0 -spezifischen Filters $F_{\vartheta_0}(f, T)$ multipliziert wird. $F_{\vartheta_0}(f, T)$ entsteht durch spektrale Glättung von $Z(\vartheta - \vartheta_0)$. Diese Glättung erfolgt z.B. durch spektrale Diffusion:

30

$$F_{\vartheta_0}(f, T) = Z(\vartheta(f, T) - \vartheta_0) + D\Delta_f^2 Z(\vartheta(f, T) - \vartheta_0).$$

Dabei bezeichnet D die Diffusionskonstante, welche ein frei wählbarer Parameter größer oder gleich null ist. Der diskrete Diffusionsoperator Δ_f^2 ist eine Abkürzung für

$$\Delta_f^2 Z(\vartheta(f,T)-\vartheta_0) = (Z(\vartheta(f-f_A/a),T)-\vartheta_0) - 2Z(\vartheta(f,T)-\vartheta_0) + Z(\vartheta(f+f_A/a,T)-\vartheta_0) / (f_A/a)^2.$$

Der auftretende Quotient f_A/a aus Abtastrate f_A und Anzahl a der Abtastwerte entspricht dem Abstand zweier Frequenzen im diskreten Spektrum. Durch Anwendung des so erzeugten Filters $F_{s0}(f,T)$ entsteht ein Spektrum
10 $S_{s0}(f,T) = F_{s0}(f,T)M(f,T)$, welches durch inverse Fouriertransformation in das Zeitsignal $s_{s0}(t)$ übergeht.

Das von einem Signalgenerator (40) auszugebene Signal $s_{s0}(t)$ entspricht dem Schallsignal innerhalb desjenigen Raumbereichs, der durch die
15 Zuordnungsfunktion $Z(\vartheta)$ und den Winkel ϑ_0 definiert ist. Der Einfachheit halber wird in der gewählten Nomenklatur für verschiedene Signalgeneratoren nur von einer Zuordnungsfunktion $Z(\vartheta)$ ausgegangen, verschiedene Signalgeneratoren benutzen lediglich verschiedene Winkel ϑ_0 . Praktisch spricht natürlich nichts dagegen, in jedem Signalgenerator auch eine eigene Form der Zuordnungsfunktion zu wählen.
20 Die Anwendung von Zuordnungsfunktionen, welche über die Zugehörigkeit von Signalkomponenten zu verschiedenen Raumbereichen entscheiden, ist einer der zentralen Gedanken der Erfindung. Eine Zuordnungsfunktion muss eine gerade Funktion sein, geeignete Funktionen sind z.B. $Z(\vartheta) = ((1 + \cos \vartheta)/2)^n$ mit einem Parameter $n > 0$. Der Raumbereich, in welchem Signale mit weniger als 3db gedämpft
25 werden, entspricht einer Hyperbel mit Öffnungswinkel $2\gamma_{3db}$ (Fig. 3) und Scheitel bei dem Winkel ϑ_0 . Hierbei entspricht $2\gamma_{3db}$ dem Halbwertswinkel der Zuordnungsfunktion $Z(\vartheta)$ (Fig. 2), mit der angegebenen Formel für die Zuordnungsfunktion gilt $\gamma_{3db} = \arccos(2^{1/n} - 1)$. Bei diesen zweidimensionalen geometrischen Überlegungen ist zu beachten, dass der tatsächliche Bereich des
30 dreidimensionalen Raums, aus welchem mit dem beschriebenen Verfahren Schallsignale extrahiert werden, ein Rotationshyperboloid ist, der durch Rotation der beschriebene Hyperbel um die Mikrofonachse entsteht.

Selbstverständlich ist die vorliegende Erfindung nicht auf den Einsatz in Kraftfahrzeugen und Freisprecheinrichtungen beschränkt: Weitere Anwendungen sind Konferenz-Telefonanlagen, bei denen mehrere Richthyperbeln in verschiedene Raumrichtungen gelegt werden, um die Sprachsignale einzelner Personen zu extrahieren und Rückkopplungen bzw. Echo-Effekte zu vermeiden. Des weiteren lässt sich das Verfahren mit einer Kamera kombinieren, wobei die Richthyperbel stets in die gleiche Richtung blickt wie die Kamera, und so nur aus dem Bildbereich kommende Schallsignale aufgezeichnet werden. In Bildtelefonsystemen ist mit der Kamera zugleich ein Monitor verbunden, in den die Mikrofonanordnung ebenfalls eingebaut werden kann, um eine Richthyperbel senkrecht zur Monitor-Oberfläche zu generieren, denn es ist zu erwarten, dass sich der Sprecher vor dem Monitor befindet.

Eine ganz andere Klasse von Anwendungen ergibt sich, wenn man anstatt des auszugebenden Signals den ermittelten Einfallswinkel ϑ auswertet, indem man z.B. zu einem Zeitpunkt T über Frequenzen f mittelt. Ein solche $\vartheta(T)$ -Auswertung kann zu Überwachungszwecken benutzt werden, wenn innerhalb eines ansonsten ruhigen Raums die Position einer Schallquelle geortet werden soll.

Das richtige "Ausschneiden" des gewünschten Bereichs entsprechend dem zu separierenden Nutz-Schallsignal aus einem Mikrofonspektrum muss nicht, wie in Figur 6 beispielhaft gezeigt, durch Multiplikation mit einer Filterfunktion erfolgen, deren Zuordnungsfunktion den in Figur 2 gezeigten beispielhaften Verlauf hat. Jede andere Art der Verknüpfung des Mikrofonspektrums mit einer Filterfunktion ist geeignet, solange diese Filterfunktion und diese Verknüpfung dazu führen, dass Werte im Mikrofonspektrum umso stärker "gedämpft" werden, je weiter ihr zugeordneter Einfallswinkel ϑ vom bevorzugten Einfallswinkel ϑ_0 (beispielsweise der Richtung des Fahrers im Kraftfahrzeug) entfernt ist.

Bezugszeichenliste:

	10	Rechenwerk zur Durchführung der erfindungsgemäßen Verfahrensschritte
5	20	Stereo-Abtast- und -Fourier-Transformator-Einheit
	30	ϑ -Berechnungseinheit
	40	Signalgenerator
	a	Anzahl der Abtastwerte, die zu Spektren M1 bzw. M2 transformiert werden
10	d	Mikrofonabstand
	D	Diffusionskonstante, wählbarer Parameter größer oder gleich Null
	Δ_f^2	Diffusionsoperator
	f	Frequenz
	f_A	Abtastrate
15	K1	erste Korrekturfunktion
	K2	zweite Korrekturfunktion
	m1(t)	Zeitsignal des ersten Mikrofons
	m2(t)	Zeitsignal des zweiten Mikrofons
	M1(f,T)	Spektrum zum Zeitpunkt T des ersten Mikrofonsignals
20	M2(f,T)	Spektrum zum Zeitpunkt T des zweiten Mikrofonsignals
	M(f,t)	Spektrum zum Zeitpunkt T des korrigierten Mikrofonsignals
	$s_{\vartheta_0}(t)$	erzeugtes Zeitsignal, entsprechend einem Winkel ϑ_0 der Richthyperbel
	$S_{\vartheta_0}(f,T)$	Spektrum des Signals $s_{\vartheta_0}(t)$
25	γ_{3db}	Winkel, welcher die Halbwertsbreite einer Zuordnungsfunktion Z(ϑ) bestimmt
	φ	Phasenwinkel des komplexen Quotienten M2/M1
	$\vartheta(f,T)$	Einfallswinkel einer Signalkomponente, gemessen von der Mikrofonachse
30	ϑ_0	Winkel des Scheitels einer Richthyperbel, Parameter in Z($\vartheta-\vartheta_0$)
	x, x1	Zwischengrößen bei der ϑ -Berechnung
	t	Zeitbasis der Signalabtastung

T Zeitbasis der Spektrumerzeugung
Z(9) Zuordnungsfunktion

Ansprüche

1. Verfahren zur Separierung von Schallsignalen von einer Mehrzahl von Schallquellen (S_1 , S_2), umfassend die Schritte:

- Anordnen von zwei Mikrofonen (MIK1, MIK2) in einem vorbestimmten Abstand (d) zueinander;
- Erfassen der Schallsignale mit beiden Mikrofonen (MIK1, MIK2) und Erzeugen zugeordneter Mikrofonensignale (m_1 , m_2); und
- Separieren des Schallsignals einer der Schallquellen (S_1) von den Schallsignalen der anderen Schallquellen (S_2) basierend auf den Mikrofonensignalen (m_1 , m_2),

dadurch gekennzeichnet, dass der Schritt des Separierens ferner die Schritte umfaßt:

- Fourier-Transformieren der Mikrofonensignale zur Ermittlung ihrer Frequenzspektren (M_1 , M_2);
- Bestimmen der Phasendifferenz (φ) zwischen den beiden Mikrofonensignalen (m_1 , m_2) für jede Frequenzkomponente ihrer Frequenzspektren (M_1 , M_2);
- Bestimmen des Einfallswinkels (ϑ) jedes einer Frequenz der Frequenzspektren (M_1 , M_2) zugeordneten Schallsignals basierend auf der Phasendifferenz (φ) und der Frequenz;
- Erzeugen eines Signalspektrums (S) eines auszugebenden Signals durch Verknüpfen eines der beiden Frequenzspektren (M_1 , M_2) mit einer Filterfunktion (F_{ϑ_0}), die derart gewählt ist, dass Schallsignale aus einer Umgebung (γ_{3dB}) um einen bevorzugten Einfallswinkel (ϑ_0) herum relativ zu Schallsignalen von außerhalb dieser Umgebung (γ_{3dB}) verstärkt werden; und
- inverses Fourier-Transformieren des derart erzeugten Signalspektrums.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die Filterfunktion (F_{ϑ_0}) ϑ -abhängig ist und unter Variation von ϑ ein Maximum bei dem

bevorzugten Einfallswinkel (ϑ_0) aufweist, und die Verknüpfung der Filterfunktion (F_{ϑ_0}) mit einem der beiden Frequenzspektren eine Multiplikation derselben umfasst.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, dass die
5 Filterfunktion (F_{ϑ_0}) die Form aufweist:

$$F_{\vartheta_0}(f, T) = Z(\vartheta - \vartheta_0) + D \Delta_f^2 Z(\vartheta - \vartheta_0)$$

wobei

f die jeweilige Frequenz

T der Zeitpunkt der Ermittlung der Frequenzspektren (M1, M2)

10 $Z(\vartheta - \vartheta_0)$ eine Zuordnungsfunktion mit Maximum bei ϑ_0

$D \geq 0$ eine Diffusionskonstante, und

Δ_f^2 ein diskreter Diffusionsoperator ist.

4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, dass die
15 Zuordnungsfunktion (Z) die Form aufweist:

$$Z(\vartheta - \vartheta_0) = \left(\frac{1 + \cos(\vartheta - \vartheta_0)}{2} \right)^n$$

wobei $n > 0$ ist.

5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch
20 gekennzeichnet, dass die Bestimmung des Einfallswinkels ϑ über die Beziehung

$$\vartheta = \arccos(x(f, T))$$

erfolgt mit

$$x(f, T) = \varphi c / 2\pi f d$$

wobei

25 φ die Phasendifferenz zwischen den beiden Mikrofonsignalkomponenten
(m1, m2)

c die Schallgeschwindigkeit

f die Frequenz der Schallsignalkomponente und

d der vorbestimmte Abstand der beiden Mikrofone (MIK1, MIK2) ist.

6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, dass es ferner den Schritt umfasst:

Begrenzen des Werts von $x(f,T)$ auf das Intervall $[-1, 1]$.

7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, dass es ferner den Schritt umfasst:

Reduzieren von Signalkomponenten, für die der Wert von $x(f,T)$ vor der Begrenzung außerhalb des Intervalls $[-1, 1]$ lag.

8. Vorrichtung zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 7, umfassend:

- zwei Mikrofone (MIK1, MIK2);
- eine an die Mikrofone angeschlossene Abtast- und Fourier-Transformationseinheit (20) zum Diskretisieren, Digitalisieren und Fourier-Transformieren der Mikrofonsignale (m_1, m_2);
- eine an die Abtast- und Fourier-Transformationseinheit (20) angeschlossene Berechnungseinheit (30) zur Berechnung des Einfallswinkels (ϑ) jeder Schallsignalkomponente; und
- wenigstens einen an die Berechnungseinheit (30) angeschlossenen Signalgenerator (40) zur Ausgabe des separierten Schallsignals.

9. Vorrichtung nach Anspruch 8, dadurch gekennzeichnet, dass der Abstand (d) zwischen den Mikrofonen der Relation genügt:

$$d < c/4f_A$$

wobei c die Schallgeschwindigkeit und f_A die Abtastfrequenz der Abtast- und Fourier-Transformationseinheit (20) ist.

10. Vorrichtung nach Anspruch 8 oder 9, dadurch gekennzeichnet, dass die Vorrichtung für jede zu separierende Schallquelle (S_1, S_2) einen Signalgenerator (40) umfaßt.

Zusammenfassung

Bei einem Verfahren zur Separierung von Schallsignalen von einer Mehrzahl von Schallquellen, umfassend die Schritte

- Anordnen von zwei Mikrofonen (MIK1, MIK2) in einem vorbestimmten Abstand (d) zueinander;
- Erfassen der Schallsignale mit beiden Mikrofonen (MIK1, MIK2) und Erzeugen zugeordneter Mikrofonssignale ($m1$, $m2$); und
- Separieren des Schallsignals einer der Schallquellen ($S1$) von den Schallsignalen der anderen Schallquellen ($S2$) basierend auf den Mikrofonausgangssignalen ($m1$, $m2$),

wird vorgeschlagen, dass der Schritt des Separierens umfaßt:

- Fourier-Transformieren der Mikrofonausgangssignale zur Ermittlung ihrer Frequenzspektren ($M1$, $M2$);
- Bestimmen der Phasendifferenz (φ) zwischen den beiden Mikrofonausgangssignalen ($m1$, $m2$) für jede Frequenzkomponente ihrer Frequenzspektren ($M1$, $M2$);
- Bestimmen des Einfallswinkels (ϑ) jedes einer Frequenz der Frequenzspektren ($M1$, $M2$) zugeordneten Schallsignals basierend auf dem relativen Phasenwinkel (φ) und der Frequenz;
- Erzeugen eines Signalspektrums (S) eines auszugebenden Signals durch Verknüpfen eines der beiden Frequenzspektren ($M1$, $M2$) mit einer Filterfunktion (F_{ϑ_0}), die derart gewählt ist, dass Schallsignale aus einer Umgebung (γ_{3dB}) um einen bevorzugten Einfallswinkel (ϑ_0) herum relativ zu Schallsignalen von außerhalb dieser Umgebung (γ_{3dB}) verstärkt werden; und
- inverses Fourier-Transformieren des derart erzeugten Signalspektrums.

(Fig. 1)

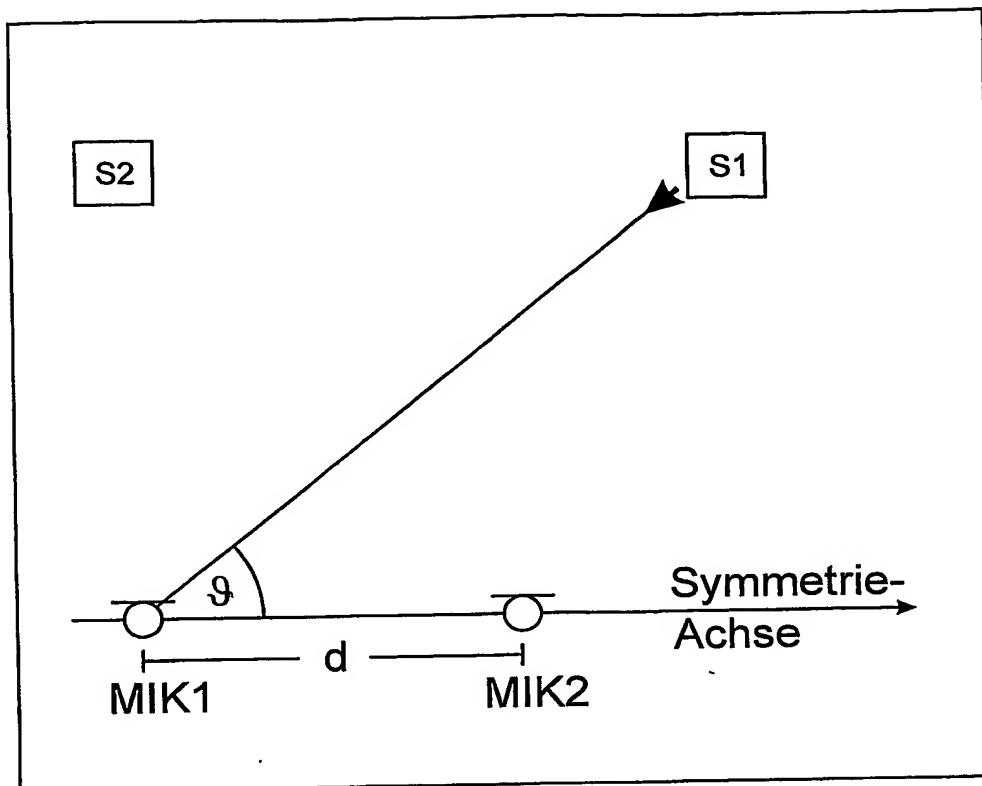


Fig. 1

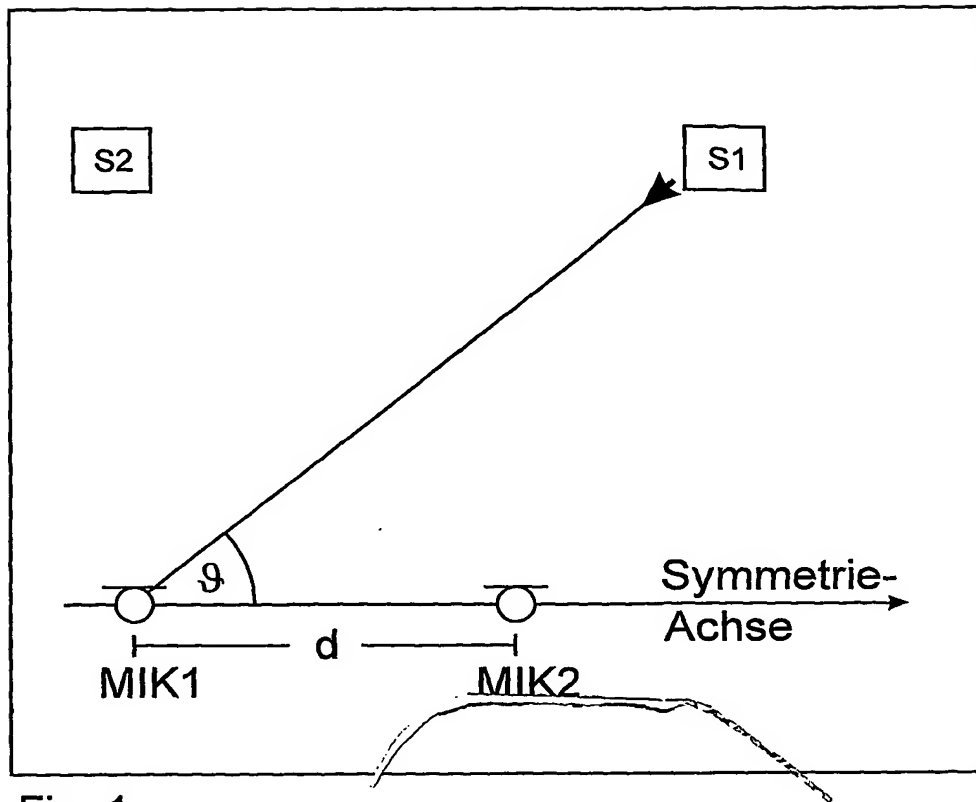


Fig. 1

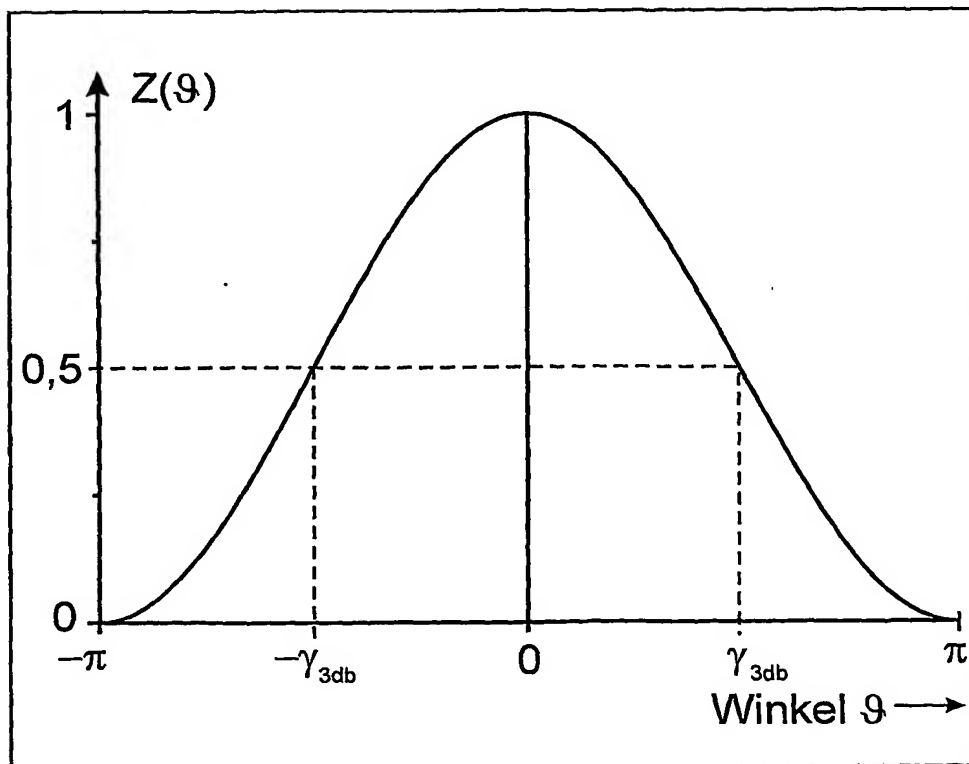


Fig. 2

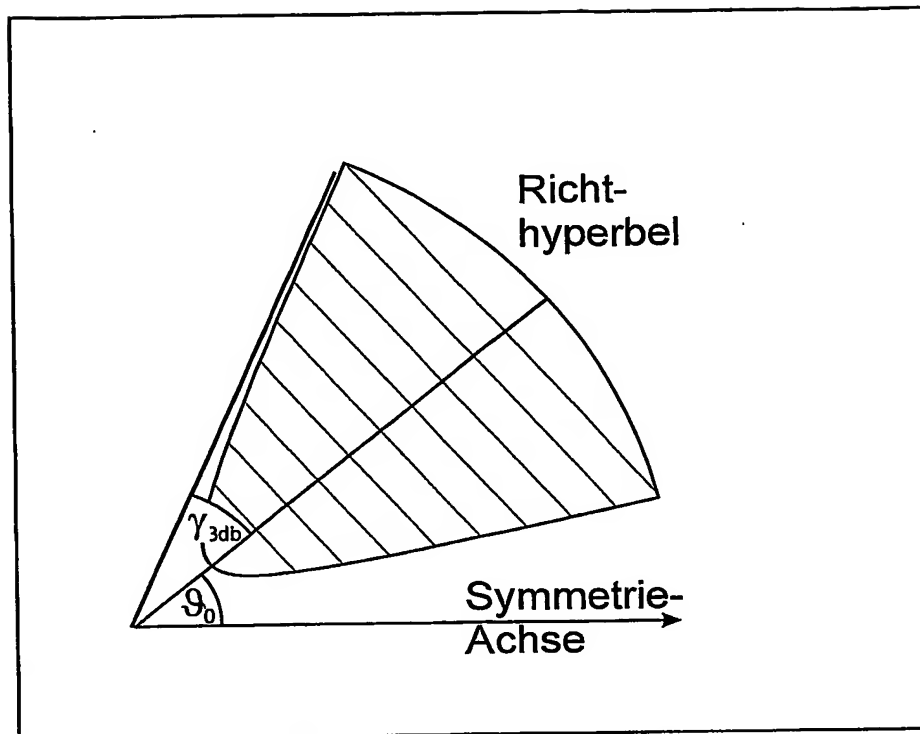


Fig. 3

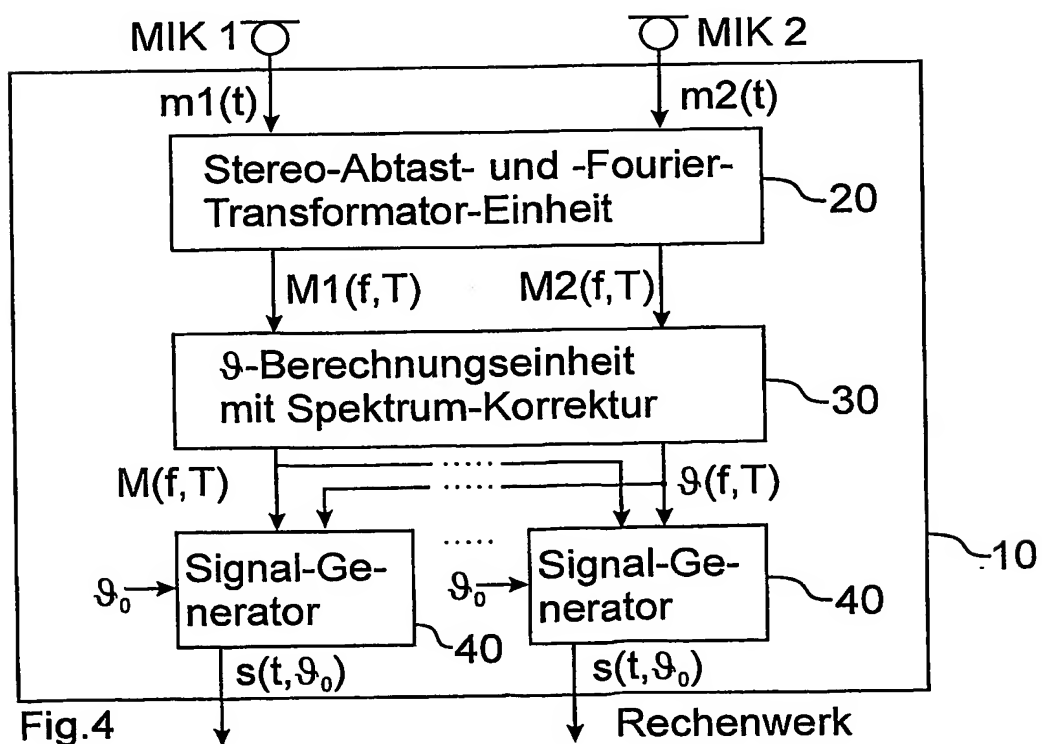
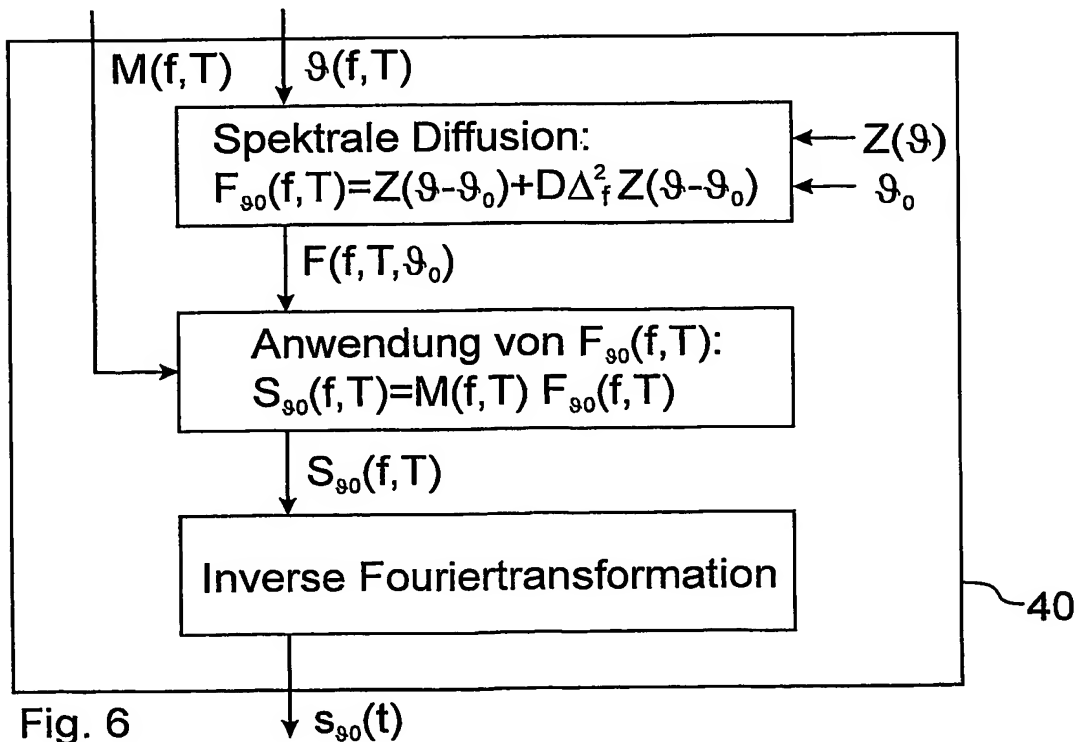
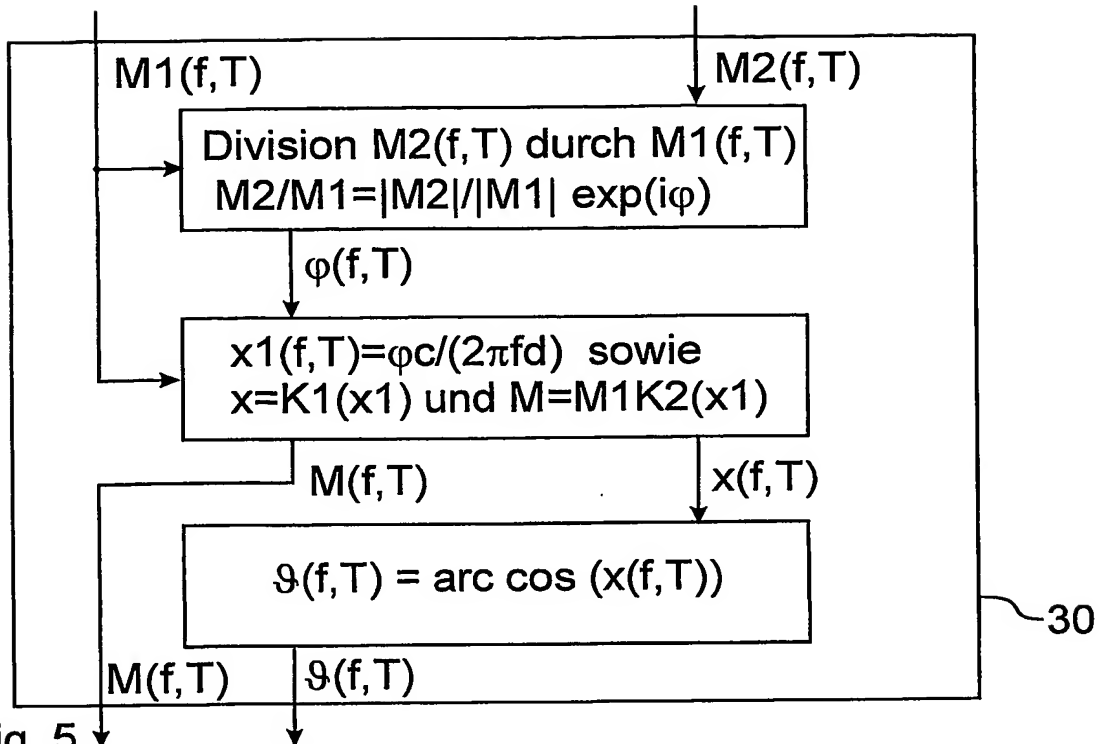


Fig. 4



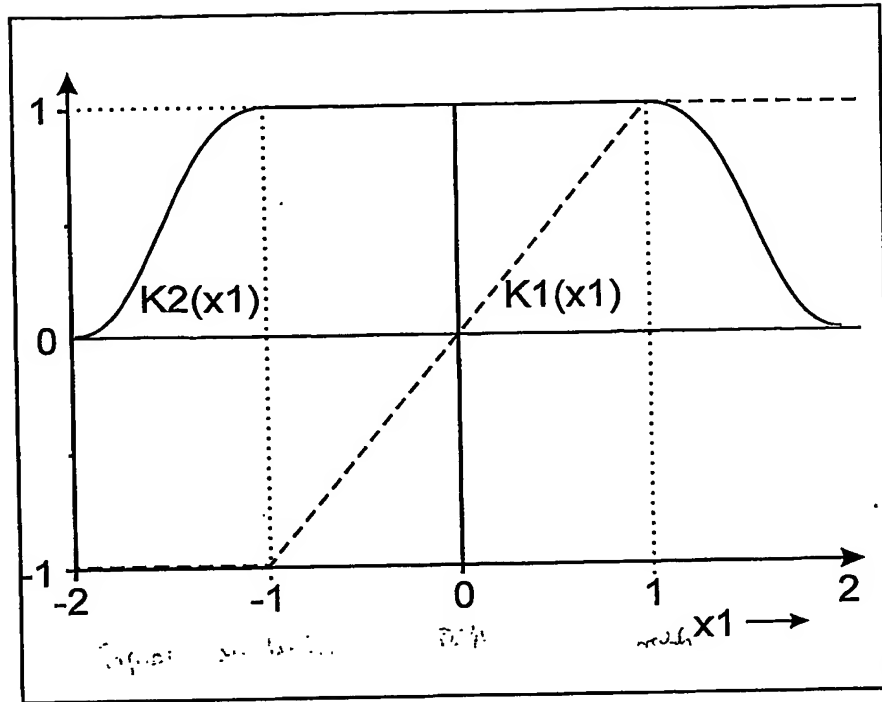


Fig. 7

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP05/050386

International filing date: 31 January 2005 (31.01.2005)

Document type: Certified copy of priority document

Document details: Country/Office: DE
Number: 10 2004 005 998.5
Filing date: 06 February 2004 (06.02.2004)

Date of receipt at the International Bureau: 28 February 2005 (28.02.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse